CGI2019 manuscript No.
(will be inserted by the editor)

Stereoscopic Image Stitching with Rectangular Boundaries

Yun Zhang · Yu-Kun Lai · Fang-Lue Zhang

Abstract This paper proposes a novel algorithm for stereoscopic image stitching, which aims to produce stereoscopic panoramas with rectangular boundaries. As a result, it provides wider field of view and better viewing experience for users. To achieve this, we formulate stereoscopic image stitching and boundary rectangling in a global optimization framework that simultaneously handles feature alignment, disparity consistency and boundary regularity. Given two (or more) stereoscopic images with overlapping content, each containing two views (for left and right eyes), we represent each view using a mesh and our algorithm contains three main steps: We first perform a global optimization to stitch all the left views and right views simultaneously, which ensures feature alignment and disparity consistency. Then, with the optimized vertices in each view, we extract the irregular boundary in the stereoscopic panorama, by performing polygon Boolean operations in left and right views, and construct the rectangular boundary constraints. Finally, through a global energy optimization, we warp left and right views according to feature alignment, disparity consistency and rectangular boundary constraints. To show the effectiveness of our method, we further extend our method to disparity adjustment and stereoscopic stitching with large horizon. Experimental results show that our method can produce visually pleasing stereoscopic panoramas without noticeable distortion or visual fatigue, thus resulting in satisfactory 3D viewing experience.

Keywords stereoscopic image stitching · rectangular boundaries · global optimization · rectangling · disparity consistency

Yun Zhang ✉
Institute of Zhejiang Radio and TV Technology, Communication University of Zhejiang, Hangzhou 310018, China.
zhangyun_zju@zju.edu.cn

Yu-Kun Lai
School of Computer Science and Informatics, Cardiff University, Cardiff CF24 3AA, UK.
Yukun.Lai@cs.cardiff.ac.uk

Fang-Lue Zhang
School of Engineering and Computer Science, Victoria University of Wellington, Wellington 6012, New Zealand.
fanglue.zhang@ecs.vuw.ac.nz

1 INTRODUCTION

As VR/AR is becoming popular, more and more people enjoy watching stereoscopic media for immersive viewing experience. Stereoscopic media is getting more popular and accessible, and the demand for manipulating stereoscopic images and videos to suit the needs of content creation has risen in many application scenarios. As the most common way to combine stereoscopic media from different cameras to generate a wider field of view, there is an apparent need for practical and reliable algorithms for stereoscopic image/video stitching. Furthermore, the stitching method can utilize images and videos captured by common hand-held cameras to provide more usable visual content.

Unlike conventional 2D visual media, stereoscopic stitching methods need to consider the preservation of depth information. Although 2D image/video stitching techniques are well studied and have been successfully integrated in digital cameras, they cannot be directly applied to stereoscopic content. The stereoscopic panorama, generated by directly applying a 2D stitching method to left and right views independently, usually has a serious disparity inconsistency, which would cause uncomfortable viewing experience when watching this region. Thus, researchers put efforts on designing

stereoscopic stitching methods that can simultaneously align features and preserve the disparity consistency. Zhang et al. [30] were the first to propose stereoscopic image stitching method to easily create stereoscopic panoramas. Although successfully keeping the disparity consistency, they did not combine the stitching process in left and right views in a unified optimization framework, and therefore cannot ensure an optimal solution with minimum distortions. Recently, Yan et al. [26] and Wang et al. [24] proposed hybrid warping and natural shape preserving methods to further improve the stitching performance. However, their methods cannot ensure the global optimality and boundary regularity.

Considering the application scenario of stereoscopic stitching, one of the major purposes of generating stereoscopic panoramas is to enlarge the viewing area. Therefore, besides the requirement of preserving the shape and disparity, we also prefer a larger visible area when watching the stitched stereoscopic panorama using common display devices, which have rectangular shapes. Thus, it would be preferable if the final panorama could have a more regular shape. However, previous methods like [26] usually generate panoramas with irregular boundaries, and the final panoramas after cropping may lose much information, which degrades the wide angle viewing effects.

Inspired by He et al. [10], who proposed a warping-based approach to rectangling 2D panoramas, we propose to stitch stereoscopic images to produce panoramas with rectangular boundaries. Our approach takes into account the characteristics of stereoscopic images, and formulates the problem as a global optimization that simultaneously optimizes stitching and rectangling. To preserve the shape during stitching as much as possible, we also propose a global optimization to simultaneously deform left and right views while preserving the shape and depth consistency. In our method, we first stitch the left and right views by optimizing feature alignment, disparity consistency, shape preservation and global similarity to form the initial state. Then we extract the irregular boundaries of stitched panoramic image pair by polygon Boolean operations in each stitched view. With the extracted vertices on the irregular boundaries, we set up regular boundary constraints, and perform an optimization that aligns stereoscopic images while satisfying these additional boundary constraints. Finally, we obtain the stitched stereoscopic panorama with a rectangular boundary by warping and blending.

The main contributions of this paper are summarized as follows:

- We are the first to propose a method for stereoscopic image stitching with rectangular boundaries, which can provide wider field of view in the generated stereoscopic panorama, compared with existing approaches.
- We formulate a global optimization to simultaneously stitch left and right views, while ensuring disparity consistency and shape preservation, outperforming state-of-the-art methods.

2 RELATED WORK

2.1 Image stitching

2D image stitching has been extensively studied in recent years, and the technology has been successfully integrated in digital cameras for panoramic image generation. In [21], Szeliski provided a tutorial on image alignment and stitching, and discussed open research problems in this area. Earlier methods mainly focused on improving the accuracy of image alignment. For global alignment, Brown et al. [1] and Gao et al. [7] proposed to use single and dual homography to warp images. However, their global warping approaches are not effective for scenes with large parallax or non-local perspective variations. To solve those problems, Zaragoza et al. [28] proposed as-projective-as-possible (APAP) warping based on moving Direct Linear Transformation (DLT), which applies several local parametric warps for more accurate alignment. Given the good performance in local feature alignment, APAP [28] is widely used in recent work [13]. Chang et al. [3] combined projective and similarity transformations for perspective preservation of each image. However, this might be too flexible and result in distortions when the spatial relationship between images is just a 2D transformation. Lin et al. [13] proposed as-natural-as-possible stitching, which combines the local homography and global similarity transformations to generate more natural panoramas. Their method can avoid unnatural rotation and mitigate perspective distortions in non-overlapping regions. However, the global similarity estimation may still introduce local distortions for some objects with highly structured geometric features. To achieve more natural appearance after warping, Chen et al. [4] proposed a stronger global similarity prior, which ensures similarity transform between each stitched image pair. Compared with [13], [4] can provide correct scales and views for better global transformation, thus can produce more natural panoramas with wider field of view. Observing that overlapping regions do not need to be fully aligned, Zhang et al. [29] proposed a local stitching method to cope with parallax, but it may fail on images with large parallax. Lin et al. [14] combined local alignment and

seam estimation by adaptive feature weighting to handle large parallax. However, their method may also produce poor results in regions containing rich structures but few matched features.

Although image stitching has been extensively studied, previous stitching methods produce panoramas with irregular boundaries for images captured by free moving hand-held cameras. However, users prefer to view images with regular boundaries, and images are usually displayed in a rectangular window. Thus, to obtain stitched results with regular boundaries, He et al. [10] first proposed panoramic image rectangling, which takes a stitched panorama as input and aims to produce rectangular images by content-aware warping. Although effective in many examples, their method treats stitching and rectangling as two steps, which cannot ensure an optimal solution and may lead to unnecessary distortions. To address this, Zhang et al. [31] incorporated rectangling into the image stitching framework to improve results, and further proposed a piecewise rectangling method to avoid excessive distortions for panoramic scenes with substantial missing content. Inspired by [10,31], we propose the first approach to stereoscopic image stitching with rectangular boundaries, which enhances the wide field viewing experience.

2.2 Stereoscopic image and video editing

In general, trivially extending 2D image/video editing methods to each view of a stereoscopic image pair will introduce disparity inconsistency and visual fatigue. For better 3D viewing experience, stereoscopic editing should firstly consider the disparities between two views.

In recent years, many researchers focused on disparity editing to improve the visual comfortableness of users. Lang et al. [11] proposed a number of disparity mapping operators to control and retarget the depths of a stereoscopic scene in a nonlinear fashion. These operators have been widely used in depth adjustment in stereoscopic images and videos for better perceptual quality. Chang et al. [2] and Liu et al. [17] proposed the content-aware stereoscopic image and video resizing methods respectively, which can adapt images and videos to different target display devices, and adjust the perceived depth in a visual comfort zone. Lee et al. [12] proposed a nonlinear disparity remapping system to make the depth manipulation more intuitive and effective. Wang et al. [25] presented a unified stereoscopic video disparity adjustment framework to automatically improve the 3D viewing experience.

How to manipulate the visual content in stereoscopic media as conventional 2D images and videos is another active topic. For interactive stereoscopic image composition, Tong et al. [23] proposed depth-consistent stereoscopic composition to interactively blend a 2D image into a stereoscopic image. Du et al. [5] extended the 2D perspective editing to stereoscopic images. Given a target perspective, their method can construct correspondent constraints and obtain the stereoscopic result by warping-based optimization. Mu et al. [20] proposed a patch-based method for stereo-consistent image completion, where the depth and image content in the missing region are filled simultaneously. Tang et al. [22] exploited depth information in stereoscopic images for more accurate saliency detection, and achieved good performance in some challenging examples. More recently, Yan et al. [27] utilized the saliency information to avoid undesirable distortion in important objects in their content-aware mesh warping model for depth adjustment.

Although 2D image stitching has been extensively studied, it cannot be directly applied to stereoscopic images. Zhang et al. [30] proposed the first method to stitch stereoscopic images casually taken by a stereo camera, and successfully handled parallax, cross-view consistency and disparity consistency in stereoscopic stitching. However, the stitching of two views is not treated as a global optimization, which cannot ensure the optimal alignment and warping, thus may introduce noticeable distortions. Yan et al. [26] proposed stereoscopic stitching based on a combination of projective and content-preserving warping. Compared with [30], their stitching is more robust and effective. However, they fix the first stereoscopic image pair as a reference, and other images may be severely distorted when aligning with the reference. Wang et al. [24] combined a constrained projective and a shape-preserving warp to make the stitching as natural as possible. Although better than previous methods [30,26], their method does not consider the boundary regularity constraint, thus may lose much information after cropping. In this paper, we aim to construct a global optimization to simultaneously stitch left and right views, while ensuring the disparity consistency, least local distortions, and boundary regularity.

3 Overview

Fig. 1 shows the pipeline of our stereoscopic stitching method. The input to our approach is two (or more) stereoscopic images with partial overlaps in their left and right views. For simplicity, we use two stereoscopic images to describe the algorithm, and each image contains a left view and a right view.

To ensure the efficiency, we represent each view using a mesh. We first formulate a global optimization

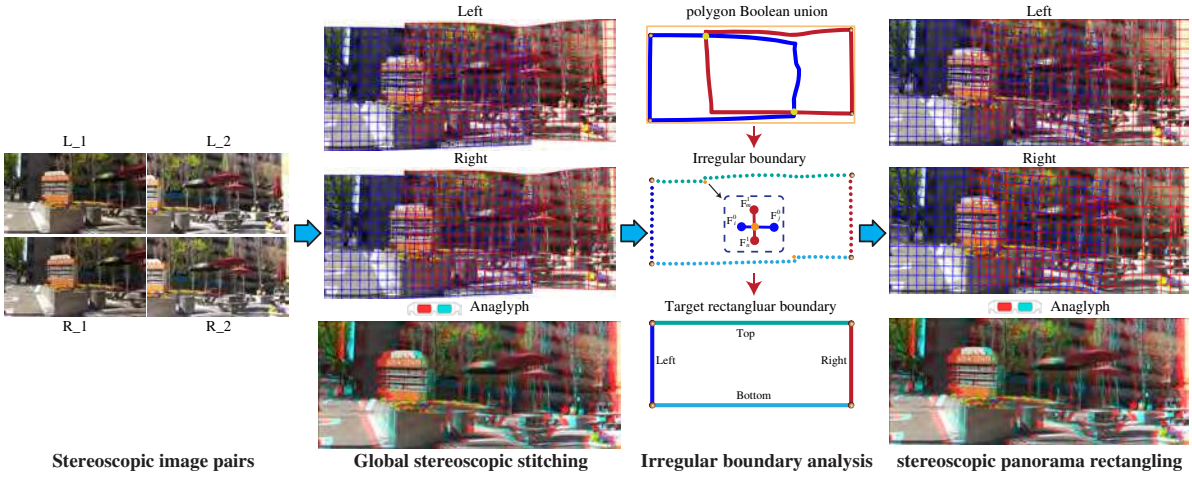


Fig. 1 Pipeline of our stereoscopic stitching. The input includes two (or more) stereoscopic images each containing a pair of (left and right) views. We firstly perform an initial global optimization for stereoscopic stitching without boundary constraints. Then we extract the irregular boundaries of the initial panorama, and construct rectangular boundary constraints. Finally we perform global optimization with the additional boundary constraints to obtain the rectangular stereoscopic panorama while ensuring feature preservation and perceptual comfortableness.

that simultaneously aligns left views and right views with constraints on feature alignment, disparity consistency, shape preserving and global similarity transform, and obtain the initial panorama by warping. Then, with the initially warped mesh on each view, we extract the irregular boundary of the initial panorama by polygon Boolean operations, and set up the target rectangular boundaries for left and right views. Finally, we perform a global optimization by adding the rectangular boundaries as constraints, and finally obtain the rectangular stereoscopic panorama by warping and blending. Our method performs stitching on left and right views using a unified global optimization framework, which can ensure seamless stitching, with less distortions and improved disparity consistency.

4 Global stereoscopic stitching

Previous methods [30,26] do not stitch left and right views by a global optimization, and thus cannot ensure the optimal disparity consistency and global structure preservation. In this section, we propose to stitch stereoscopic images in a global optimization framework, which also provides the initial results for further panorama rectangling.

Similar to previous stitching methods, we place a mesh on each view of input stereoscopic images, and the stereoscopic stitching result is obtained by warping-based optimization. Let N be the number of input stereoscopic image pairs to be stitched, V and $\hat{V}$ are mesh vertex positions before and after warping. We propose an energy involving the following terms: feature alignment of each view E_f, local shape similarity E_s, global similarity E_g, and disparity consistency E_d, and details are as follows.

Feature Alignment. Feature alignment is used to match each view (left or right) of stereoscopic images with the same view of other stereoscopic images within overlapped regions. Similar to [4], we use APAP [28] due to its high quality for image feature matching. It is defined as follows.

$$E_f(\hat{V}) = \sum_{d=0}^{1} \sum_{(i,j)\in\mathcal{G}_d} \sum_{f_k^{ij}\in M_d^{ij}} ||\hat{v}(f_{k,i}^{ij}) - \hat{v}(f_{k,j}^{ij})||^2, \quad (1)$$

where $\mathcal{G}_d$ is the image matching graph of left ($d=0$) and right ($d=1$) views, containing image pairs of the same view with overlaps. M_d^{ij} refers to all the matched features between images i and j of view d. In practice, the points $f_{k,i}^{ij}$ and $f_{k,j}^{ij}$ of a matched feature pair f_k^{ij} do not usually coincide with mesh vertices, so we represent each matched feature point $f_{k,i}^{ij}$ (similar for $f_{k,j}^{ij}$) as a bilinear interpolation of vertices in the mesh quad that contains the feature point. The interpolation $\hat{v}(f_{k,i}^{ij})$ is represented as $\hat{\mathbf{V}}_{mn}^{i} \cdot \mathbf{\Omega}_{mn}^{i}$, where $\hat{\mathbf{V}}_{mn}^{i} = [\hat{V}_{m,n}^{i}, \hat{V}_{m+1,n}^{i}, \hat{V}_{m+1,n+1}^{i}, \hat{V}_{m,n+1}^{i}]$ includes the four vertices of the mesh quad that contains the feature point, and $\mathbf{\Omega}_{mn}^{i} = [\omega_{m,n}^{i}, \omega_{m+1,n}^{i}, \omega_{m+1,n+1}^{i}, \omega_{m,n+1}^{i}]$ are the interpolation weights. $\hat{v}(f_{k,j}^{ij})$ is similarly defined.

Local shape similarity. Similar to most warping based stitching, we also need to preserve the shape of each mesh quad in stereoscopic stitching. We split each mesh quad into two triangles, and use the shape preserving term defined in [16] to penalize distortions of

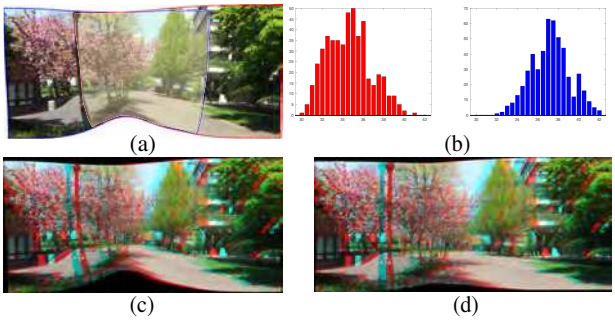


Fig. 2 Content-aware disparity constraint. (a) overlapped region of left stitching result, (b) disparity histogram comparison of two stereoscopic images in the overlapping region, (c) stitching result obtained by original disparity, (d) stitching result obtained by content-aware disparity.

triangle shapes. The energy is defined as follows.

$$E_s(\hat{V}) = \sum_{d=0}^{1}\sum_{i=1}^{N}\sum_{\hat{V}_j^i \in \hat{\mathcal{V}}_d^i} ||\hat{V}_j^i - \hat{V}_{j_1}^i - \xi \mathbf{R}(\hat{V}_{j_0}^i - \hat{V}_{j_1}^i)||^2$$
$$where \quad \mathbf{R} = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}, \quad \xi = ||V_j^i - V_{j_1}^i||/||V_{j_0}^i - V_{j_1}^i||, \tag{2}$$

where d indicates left ($d = 0$) and right ($d = 1$) views of each stereoscopic image pair. $\hat{V}_j^i$, $\hat{V}_{j_0}^i$, $\hat{V}_{j_1}^i$ are vertices of a triangle in each mesh quad after warping, and V_j^i, $V_{j_0}^i$, $V_{j_1}^i$ are corresponding vertices before warping. ξ is the scaling parameter estimated from the initial mesh. Since each triangle is right-angled before warping, we set $\mathbf{R}$ a transformation matrix with 90° rotation.

Global similarity. To make the stitching result as natural as possible, we use the global similarity term defined in [4] to constrain the rotation and scale of each stitched image according to the reference image (chosen as the first image) in each view. The energy term is define as follows.

$$E_g(\hat{V}) = \sum_{d=0}^{1}\sum_{i=2}^{N}\sum_{e_j^i \in E_i^d} \omega(e_j^i)[||c(e_j^i) - s_i \cos\theta_i||^2 + ||s(e_j^i) - s_i \sin\theta_i||^2], \tag{3}$$

where s_i, θ_i ($i = 2 \cdots N$) refer to the desired scale and rotation of each warped image in left ($d = 0$) and right ($d = 1$) views w.r.t. the corresponding views in the reference image I_1. $\omega(e_j^i)$ is the weight to determine the importance of edges for global similarity. In the overlapping regions, we need to focus on feature alignment, thus the weight should be smaller, where for other regions, larger values are given for better preserving global similarity. See [4] for detailed definitions.

Disparity consistency. Simply extending standard image stitching to stereoscopic images may cause disparity inconsistency and visual fatigue. Actually, when we consider the disparity preserving constraint in the global stitching framework, we need to balance this constraint in a content-aware manner. See Fig. 2(c), when we simply copy the disparity of input stereoscopic images, the final result may be severely distorted. In this example, the disparities of two stereoscopic images in the overlapping region (marked in Fig. 2(a)) vary greatly, see the disparity histograms in Fig. 2(b). Thus, in the overlapping regions, the disparity and feature alignment constraints conflict. For better 3D viewing effects, we add a weight $\Psi(\cdot)$ to control the disparity consistency as follows:

$$E_d(\hat{V}) = \sum_{i=1}^{N}\sum_{p_k^i \in P^i} \Psi(p_k^i) \cdot ||\hat{v}(p_{k,l}^i) - \hat{v}(p_{k,r}^i) - D(p_k^i)||^2, \tag{4}$$

where $p_{k,l}^i$ and $p_{k,r}^i$ refer to the k^{th} matched feature points in the i^{th} image pair, and we use the interpolation of quad vertices to represent each feature point in the left view as $\hat{v}(p_{k,l}^i) = \hat{\mathbf{V}}_{mn}^{i,l} \cdot \mathbf{\Omega}_{mn}^{i,l}$, where $\hat{\mathbf{V}}_{mn}^{i,l}$ are vertices of a quad that contains $p_{k,l}^i$ in the left view of the i^{th} stereoscopic image, and $\mathbf{\Omega}_{mn}^{i,l}$ refers to the corresponding interpolation weight of each vertex. Feature points $\hat{v}(p_{k,r}^i)$ in the right view are similarly defined. $D(p_k^i)$ is the original disparity of feature p_k^i, and $\Psi(p_k^i)$ is the weight for the disparity constraint. To keep the disparities of stitched panorama in a proper range, we set I_1 as a reference image, and the disparities of I_1 are preserved by a large weight. Observing that local feature alignment and disparity constraints should be balanced in the overlapping regions, we measure the weight of disparity constraint $\Psi(p_k^i)$ of a matched feature pair p_k^i in other images $I_i (i = 2, ..., N)$ as

$$\Psi(p_k^i) = \begin{cases} \delta_i \cdot \frac{d_{\min}(q(p_k^i), \Phi^i)}{\sqrt{W_i^2 + H_i^2}}, & p_k^i \notin \Phi^i \\ \delta_i \cdot (T(\Phi^i) + 0.05)^{-\alpha}, & p_k^i \in \Phi^i \end{cases}, \tag{5}$$

where Φ^i refers to the overlapping regions of the i^{th} image, and $q(p_k^i)$ is the center of the mesh quad that contains p_k^i. For matched features in non-overlapping regions, we calculate the minimum distance between $q(p_k^i)$ and all mesh quads in Φ^i, denoted as $d_{\min}(\cdot, \cdot)$. δ_i is used to place more importance on the reference image. We set $\delta_1 = 10$ and $\delta_i = 1$ ($\forall i \neq 1$). W_i and H_i are width and height of the i^{th} image. For overlapping regions containing two images, the more similar their disparity histograms are, the larger weight is on the disparity consistency term. $T(\cdot)$ measures the distance

of disparity histogram between the images in Φ^i, and α is used to control the impact of disparity similarity on the disparity constraint. In our experiment we set $\alpha = 1$. When more than two images overlap, we select the image I_m with median disparities as reference, and $T(\cdot)$ measures the calculated between I_i and I_m. See Fig. 2(d), with our content-aware disparity constraint, the stitching result is more natural with less distortion.

With the energy terms defined above, the total energy E_1 is a linear combination of them, see Equation 6.

$$E_1(\hat{V}) = \lambda_a E_f(\hat{V}) + \lambda_s E_s(\hat{V}) + \lambda_g E_g(\hat{V}) + \lambda_d E_d(\hat{V}), \quad (6)$$

where λ_a, λ_s, λ_g, λ_d are weights to determine the importance of each term. We give more importance to disparity preservation and shape preserving for less distortion and disparity consistency, and set $\lambda_a = 1$, $\lambda_s = 2$, $\lambda_g = 1.0$, $\lambda_d = 6$ for most examples.

5 Stereoscopic Rectangling

Compared with results generated by 2D image stitching and [30,26,24], our global stitching is better in shape preserving and disparity consistency, see Fig. 5. However, our result in Fig. 5(f) still generates the panorama with an irregular boundary. Thus, we further improve our global stitching by rectangling the irregular boundary of stitched panoramas. The main idea of stereoscopic rectangling is to drag the vertices on the initial irregular boundary of the stitching result to the target rectangular boundary, then the stitched image can be warped driven by these boundary constraints.

5.1 Boundary analysis

In this section, we give details of irregular boundary extraction from the warped meshes of global stereoscopic stitching. We use the method proposed in [31] for irregular boundary extraction in the corresponding view of the warped meshes. Fig. 1 shows the irregular boundary extraction on the left view. We first extract the boundary vertices of each warped mesh, and represent them as polygons. Then we get the irregular boundary of panorama by the polygon Boolean union operation [18]. The irregular boundary is represented by boundary vertices from different meshes and their intersections. We further get the outer rectangle of the irregular boundary, and select the closest mesh corners to the outer rectangle as four corners of the irregular boundary. Then we classify irregular vertices into 4 groups with different directions (left/right/top/bottom) using neighboring corners. We also identify the intersections in each direction, and represent them using interpolation of neighboring related vertices. The target rectangular boundary is then constructed by averaging vertices in each direction, see Fig. 1 for the rectangle with different colors in four directions.

5.2 Stitching with rectangular boundary

The rectangular boundary is converted to a constraint in the final global optimization step to warp the stereoscopic images to fill the rectangle. Based on the energy terms defined in Section 4, we further add the rectangular boundary E_r and line preserving E_l constraints to formulate the optimization objective. We give detailed definition of energy terms as follows.

Rectangular boundary. We define the rectangular boundary preserving energy as

$$E_r(\hat{V}) = \sum_{d=0}^{1}\sum_{i=1}^{4}\sum_{\hat{V}_j \in \xi_i^d} ||\Gamma(\xi_i^d)[\zeta[\hat{V}_j]\hat{V}_j + (1-\zeta[\hat{V}_j]) (\vartheta[\hat{V}_j]\cdot\omega[\hat{V}_j])] - val[i]||^2, \quad (7)$$

where $\xi_i^d, i = 1,2,3,4$ refers to the set of vertices in four directions of the extracted irregular boundaries in left ($d = 0$) and right ($d = 1$) views; $val[i]$ refers to the value of each target boundary segment; $\zeta[\hat{V}_j]$ determines the type of vertices: 1 - vertices, 0 - intersections; $\Gamma(\xi_i^d) = [0 \quad 1]$ or $[1 \quad 0]$, when ξ_i^d is horizontal or vertical, which determines direction of vertices to be dragged. For the intersection points, we first detect the neighboring related vertices $\vartheta[\hat{V}_j] = [\hat{V}_m, \hat{V}_n, \hat{V}_p, \hat{V}_q]$ that produce the intersections, see the dotted rectangle in Fig. 1, then we calculate the interpolation weights of the vertices $\omega[\hat{V}_j] = [C_m, C_n, C_p, C_q]^T$, finally the intersections can be easily represented by the interpolation of corresponding mesh vertices.

Straight line preserving. To avoid obvious line distortion after warping, we also need to preserve the straight lines in stereoscopic stitching. Similar to the warping-based retangling in 2D images [10,31], we detect line segments using the method in [8], and use the line preserving term defined in [15].

$$E_l(\hat{V}) = \sum_{d=0}^{1}\sum_{i=1}^{N}\sum_{l\in L_i^d}\sum_{j=1}^{M-1} ||(1-\mu)\hat{\mathbf{V}}_{pq}^{i,l_0}\cdot\mathbf{\Omega}_{pq}^{i,l_0} + \mu\hat{\mathbf{V}}_{pq}^{i,l_m}\cdot\mathbf{\Omega}_{pq}^{i,l_m} - \hat{\mathbf{V}}_{pq}^{i,l_j}\cdot\mathbf{\Omega}_{pq}^{i,l_j}||^2, \quad (8)$$

where d indicates the left or right view, N refers to the number of all stitched left and right images, M is the number of sub-segments for each line segment in an image, and each sample point on the line segment is

represented by a bilinear interpolation of the four grid vertices. The two end points l_0 and l_m are represented as $\hat{\mathbf{V}}_{pq}^{i,l_0} \cdot \mathbf{\Omega}_{pq}^{i,l_0}$ and $\hat{\mathbf{V}}_{pq}^{i,l_m} \cdot \mathbf{\Omega}_{pq}^{i,l_m}$. The j^{th} sample point between the end points is l_j, and the weight is $\mu = j/M$.

We reconstruct the global optimization for stereoscopic stitching by adding the boundary and line preserving constraints as follows.

$$E_2(\hat{V}) = E_1 + \lambda_r E_r(\hat{V}) + \lambda_l E_l(\hat{V}), \tag{9}$$

where E_1 is defined in Equation 6, and λ_r, λ_l are weights to control the importance of energy terms. We set $\lambda_r = 10^2$ to ensure the regularity of boundaries. In our experiment, we find that the line preserving is more important than the local shape preserving, thus λ_l is set to be 15 to avoid too much distortions in straight lines.

5.2.1 Optimization

We first solve the optimization defined in Equation 6 for the initial stereoscopic image stitching. Since all energy terms are quadratic, Equation 6 can be efficiently minimized by solving a linear system. Since the resulting vertices are used for the irregular boundary analysis, thus we do not need to render the stereoscopic panorama at this step. With the extracted irregular boundary and the target rectangular boundary, we further solve the reconstructed optimization in Equation 9. Since the added rectangular boundary and line preserving terms are quadratic, Equation 9 can also be minimized by solving a linear system. For better stereoscopic viewing effects, we set the same target rectangle in left and right views. The final stereoscopic panorama is obtained by warping and blending, see the right side of Fig. 1. As shown in Fig. 5, compared with previous stereoscopic stitching methods [30,26,24], our rectangular stereoscopic stitching can preserve much more contents in a rectangular window, and provide better wide angle viewing experience.

5.2.2 Content-aware rectangling

Although our stereoscopic stitching can effectively produce panoramas with rectangular boundaries, it does not work well when there exist salient objects or structures near the irregular boundary. As shown in Fig. 3(b), due to the content missing in the structured region, there exist apparent distortions.

Thus, we further propose a content-aware rectangling method, which progressively performs rectangling while avoiding noticeable distortions. We first compute E_0 of global stitching in Equation 6 using the solved vertices. Then we iteratively set each boundary ξ_i^d of the target rectangle as the boundary constraint, and solve optimization in Equation 9. With the warped vertices, we compute $E_2(\hat{V})$ and compare it with E_0. When there are few differences between them, we accept ξ_i^d as a boundary constraint, and push it into $\hat{\xi}$. In this paper, we set the threshold $\delta = E_0/15$. Finally, using the filtered boundary constraints $\hat{\xi}$, we solve the optimization defined in Equation 9 for panorama rectangling, and obtain final stereoscopic panoramas by warping and blending. Compared with the global stitching and direct rectangling (see Figs. 3(a) and 3(b)), our content-aware rectangling (Fig. 3(d)) can make the panorama as rectangular as possible without noticeable visual distortions, which further enhances the wide angle viewing effects, and Fig. 3(c) shows the warped meshes. The *yellow* rectangle in (d) gives the suggested cropping window, which is wider than that in (a).

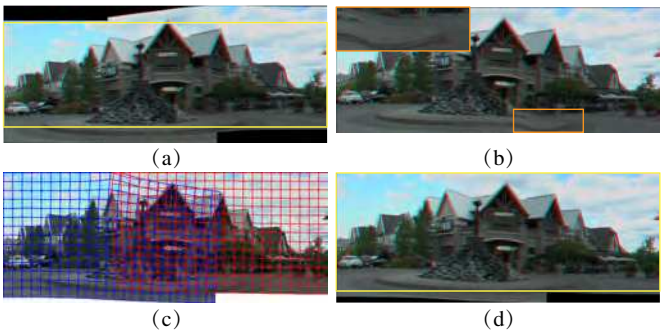


Fig. 3 Content-aware stereoscopic rectangling. (a) global stereoscopic stitching, (b) stereoscopic stitching with rectangling, (c) content-aware stereoscopic rectangling meshes, (d) content-aware stereoscopic rectangling result.

6 RESULTS AND APPLICATIONS

In this section, we give a number of experimental results of stereoscopic stitching, and applications that benefit from our method. Then performance, evaluations and limitations of our method are provided. We use the dataset from [30,26] for most examples in this paper. In the following comparisons, results of [30], provided by their authors, have been cropped by rectangular windows. In this paper, the stitching results are presented as red&cyan anaglyph images.

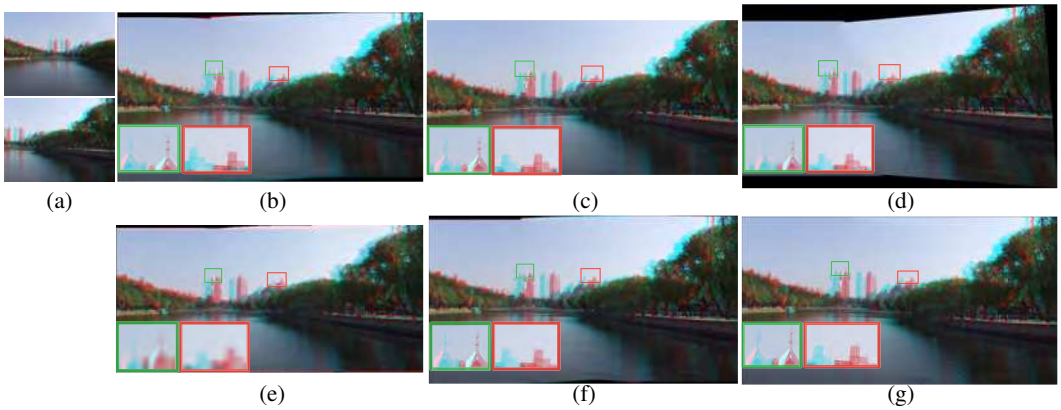
For better viewing experience, we provide high resolution input and output stereoscopic images of all examples in the supplemental material.

6.1 Results and Comparisons

As shown in Fig. 4, we first compare our method with state-of-the-art methods. The *red* rectangles in each result demonstrate that [30] introduces distortion, while

(a) (b) (c)

Fig. 4 Comparison with state-of-the-art methods. (a) result of [30] , (b) result of [24], (c) our final stitching result.



(a) (b) (c) (d)

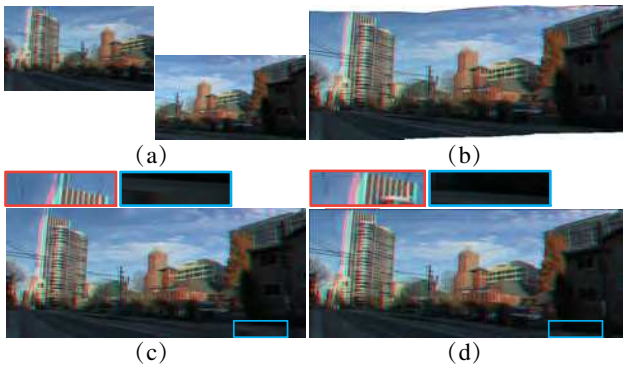
(e) (f) (g)

Fig. 5 More comparisons. (a) input stereoscopic images, (b) result obtained by separate stitching, (c) result of [30], (d) result of [26], (e) result of [24], (f) our global stitching result, (g) our stitching result with rectangular boundary.

[24] and our method succeed in preserving shapes. The zoom-in views show that both [24] and [30] cannot ensure a regular boundary in the stitched panorama, while our method works well for boundary regularity, thus can preserve more contents in the cropping window.

We further give more comparisons in Fig. 5. When stitching two views separately or stitching using the method in [24], there are obvious vertical disparities, as shown in the zoom-in view of Figs. 5(b) and 5(e)[1]. Figs. 5(c) and 5(d) are results of [30] and [26], which may lose some image contents after cropping. Fig. 5(f) is the result obtained by our global stitching, which is better that previous methods in terms of content preserving and disparity consistency. By adding the regular boundary constraint in our global optimization, the stitching result can preserve all image content in a rectangle, see Fig. 5(g).

An alternative to handling the irregular stitching results could be content-aware completion. Fig. 6 shows comparison of our stereoscopic rectangling with it. We first stitch the input stereoscopic images in Fig. 6(a) using our global stitching method. Then we use the stereoscopic completion method in [20] to complete the hole between outer rectangle and the irregular boundary of panorama (see Fig. 6(b)). The zoom-in view of Fig. 6(c) shows that the completion method fails to synthesize correct semantic content. Fig. 6(d) is the result of our method, which can successfully generate stereoscopic panorama with rectangular boundary, while preserving the disparity consistency and avoiding distortions.



(a) (b)

(c) (d)

Fig. 6 Stereoscopic completion. (a) input stereoscopic images, (b) global stitching result, (c) rectangling result obtained by completion, (d) our stereoscopic rectangling result.

We also verify the contribution of the line preserving term. Fig. 7 shows examples which demonstrate the effects of the line preserving term in stereoscopic stitching. As shown in Fig. 7(a), our global stitching does not introduce distortions. When rectangling without line preservation, the salient lines are distorted, see the zoom-in view in Fig. 7(b). For better effects, we impose the line preservation constraint, and the final result is visually pleasing, see the zoom-in view in Fig. 7(c).

More results and comparisons are shown in Fig. 8. Compared with the results in Figs. 8(a) and 8(b) generated by [30, 26], our global stitching results in Fig. 8(c)

[1] we take the result of Fig 5(e) directly from the [24] paper, as the code is not publicly available thus the zoom-in view is not very clear.

have less distortion and more regular boundary, thus provide better wide angle viewing experience. Based on the global stitching, our stitching with rectangular boundary constraint can further improve the wide angle viewing of stereoscopic panorama, see Fig. 8(d).

6.2 Applications

We further extend our method to various applications of stereoscopic image editing.

6.2.1 Disparity adjustment in panoramic images

In stereoscopic images and videos, disparities can determine the 3D feelings of human vision system (HVS), and unsuitable disparities can cause visual discomfort and fatigue. For better experiences to view the stereoscopic images, we incorporate disparity adjustment into the optimization framework of our stereoscopic stitching. In this application, we only need to modify the energy term for disparity consistency, and the new definition is as follows.

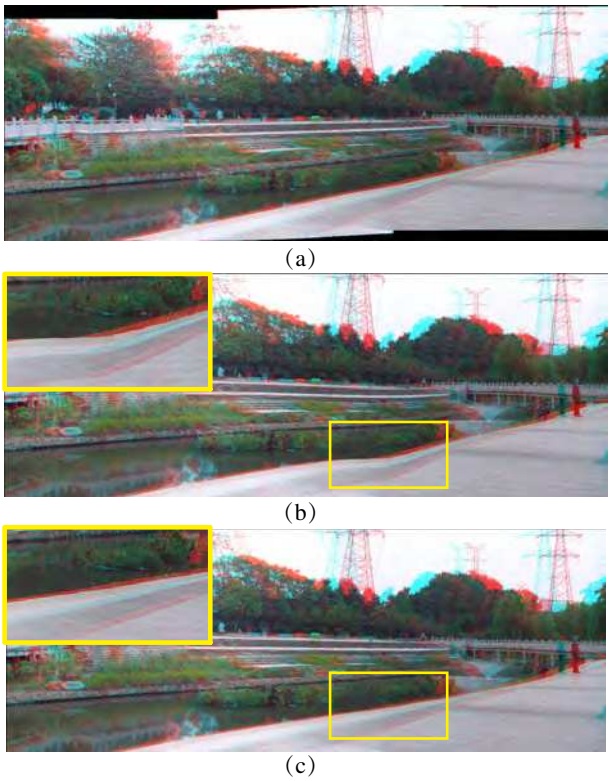


Fig. 7 Line preserving. (a) global stitching result, (b) rectangling without line preserving constraint, (c) our stitching result with line preserving constraint.

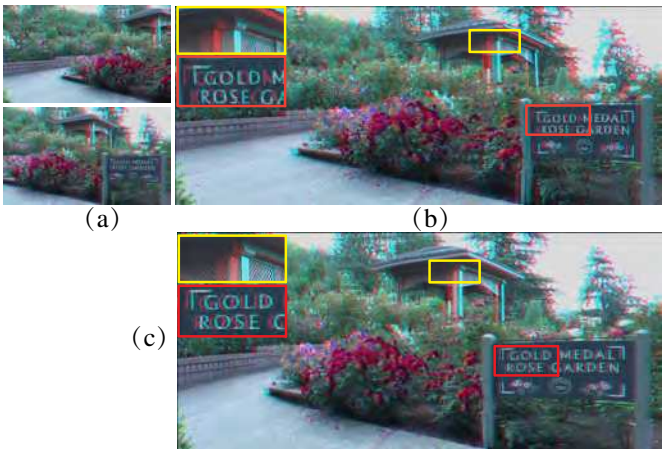


Fig. 9 Stereoscopic rectangular stitching with disparity adjustment. (a) input stereoscopic images, (b) stitching result with rectangular boundary, (c) stitching result after disparity adjustment.

$$E_d(\hat{V}) = \sum_{i=1}^{N} \sum_{p_k^i \in P^i} \Psi(p_k^i) \cdot ||\hat{v}(p_{k,l}^i) - \hat{v}(p_{k,r}^i) - M(D(p_k^i))||^2, \tag{10}$$

where $M(\cdot)$ is the disparity mapping function [11] which aims to change the disparity range of stitched panorama. Fig. 9 shows an example of disparity adjustment. For better 3D viewing, we first use the linear operator to set the disparity range in a visual comfort zone, and then further compress the disparity values by setting $M(D(p_k^i)) = \alpha \cdot D(p_k^i)$, with $\alpha \in [0, 1]$. Figs. 9(b) and 9(c) are rectangular stitching results before and after disparity adjustment, and we can obtain different 3D viewing experience without sacrificing the rectangular stitching quality.

6.2.2 Rectangling panorama with large horizon

We further apply our method to stitching stereoscopic panorama with wide-angle view. To obtain such panorama, we need to stitch a set of images with partial overlaps. Unlike [26], which stitches a number of overlapping images incrementally, we stitch a set of images using our proposed global optimization, thus can produce panorama with regular boundary, better stereoscopic effects and less distortion. Fig. 10 shows the comparison on the generated panoramas consisting of 10 overlapping stereoscopic images. Figs. 10(a) and 10(b) are results of [30] and [26] respectively. Fig. 10(c) is our stitching result without boundary constraint, which is better than Figs. 10(a) and 10(b) in terms of distortion and boundary regularity, and our result can preserve more

Fig. 8 More results and comparisons. (a) results of [30], (b) results of [26], (c) results obtained by our global stitching, (d) our stitching results with rectangular boundary.

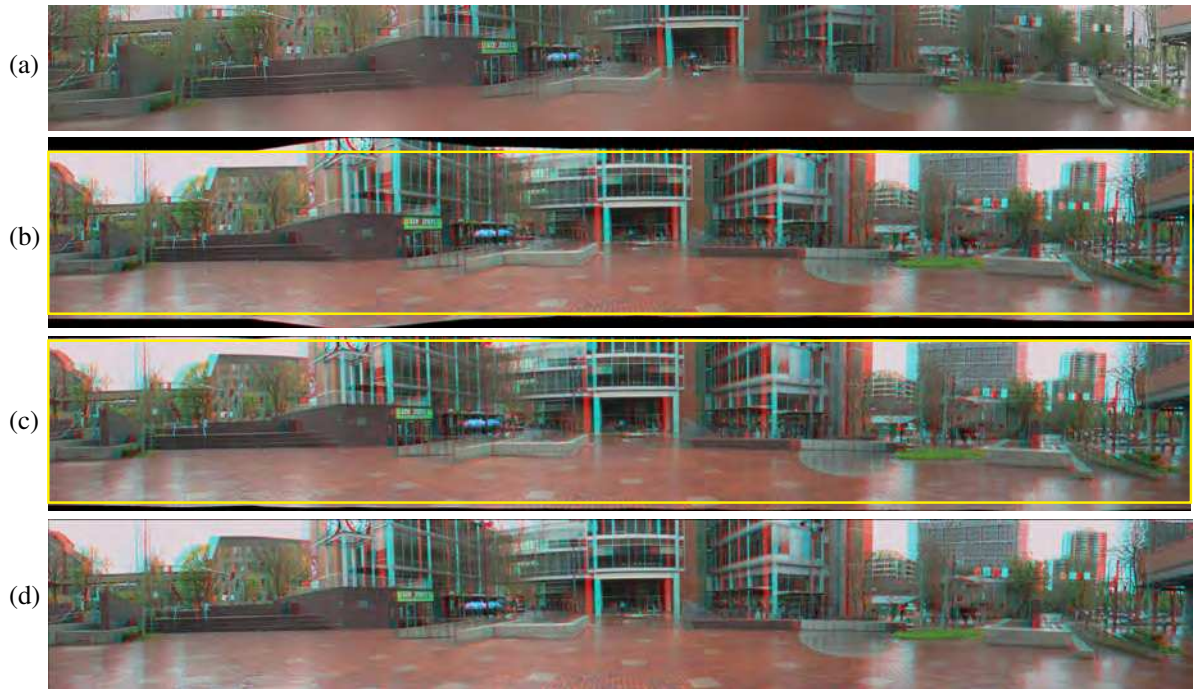


Fig. 10 Wide angle stereoscopic rectangling. Result in (a) is provide by authors of [30] and lots of content has been cropped. (b) result obtained by [26], (c) result obtained by our global stitching, (d) final stitching result with rectangular boundary.

content without noticeable distortions in a rectangular cropping window. For better wide angle viewing, we further apply rectangling to stereoscopic stitching with large horizon. The result in Fig. 10(d) shows that, our method can preserve all image contents in a rectangular window, thus significantlyh improve the wide-angle viewing experience.

6.3 Performance and Evaluations

We report performance of our method on an Intel Core i7 6700K 4.0GHz laptop with 32GB RAM. As an example, consider Fig. 1 where the input contains 2 stereoscopic images with partial overlaps, and the size of each image is 800×600, the total time cost is 6.86 seconds, which includes the initial stitching, irregular boundary extraction, stitching with rectangular boundary and texture mapping & blending. Since all energy terms are quadratic, the optimization can be efficiently solved. When stitching with the optimized boundary, we only need to construct the irregular boundary constraint, and apply the conjugate gradient method, which takes the result of last iteration as initialization, thus can solve the optimization more efficiently.

We further give quantitative evaluation, as shown in Table 1. For the five examples in Fig. 8, we compare the

Table 1 Quantitative evaluations of average vertical disparities in pixels and cropping ratio

	Chen's [4]	Zhang's [30]	Yan's [26]	Ours
#1	0.93\91%	0.61\92%	0.58\86%	0.46\100%
#2	1.67\95%	1.10\96%	0.92\94%	0.88\100%
#3	1.33\94%	0.83\93%	0.81\91%	0.57\100%
#4	1.47\92%	1.01\92%	0.95\90%	0.71\100%
#5	1.26\93%	0.74\94%	0.81\89%	0.65\100%

Table 2 User study of our method

	visual comfort	free of artifacts	wide angle
Yan [26]	$\mu(4.07)\sigma(0.57)$	$\mu(4.13)\sigma(0.50)$	$\mu(3.93)\sigma(0.57)$
Ours	$\mu(4.53)\sigma(0.49)$	$\mu(4.40)\sigma(0.49)$	$\mu(4.67)\sigma(0.47)$
P-value	0.029	0.164	0.001

average vertical disparities and cropping ratios using Chen's [4], Zhang's [30], Yan's [26] and our method. The cropping ratio is the ratio of cropped image content and the final stitching result. Compared with [26], results obtained by our method have less vertical disparities, and the cropping ratio is 100% due to our stitching with the rectangular boundary constraint.

To evaluate the visual quality of our stitching results, we asked 15 college students (7 males and 8 females) with normal stereo perception to take part in our user study. For each example, these students were asked to watch 20 stitched panoramas obtained by Yan's [26] and our method. After watching them with stereo glasses for 5∼10 minutes, they were asked to grade the following three questions: (1) Do you feel visually comfortable to watch the stereoscopic panoramas? (2) Do the stitched panoramas free of artifacts? (3) Are the stitched panoramas provide good wide angle viewing experiences? Each question was graded by an integer ranging from 0 to 5 (negative to positive). We report the average score (μ), standard deviation (σ), and P-value of the two sample t-test in Table 2. From the average scores, we find that our method has better performance than [26]. The P-value shows that the difference between the two sets of results is significant in terms of "visual comfort" and "wide angle", while not significant for the "free of artifacts". Thus, the advantage of our method is more obvious in producing stereoscopic panoramas with visual comfort and wide angle effects. In terms of "free of artifacts", our method is comparable to state-of-the-arts method.

6.4 Limitations

Our method improves stereoscopic image stitching by rectangling the irregular boundaries caused by the free movement of hand-held cameras, and works well for many cases. However, we notice that our method may fail in some cases, especially when the scene is not completely shot, and some structural features are may be distorted when the saliency detection fails. See Fig. 11 for an example, the global stitching result in Fig. 11(b) is visually pleasing, but after rectangling, the stitched rectangular panorama is severely distorted, see the zoom-in view in Fig. 11(c). This problem also exists in most warping based image editing methods.

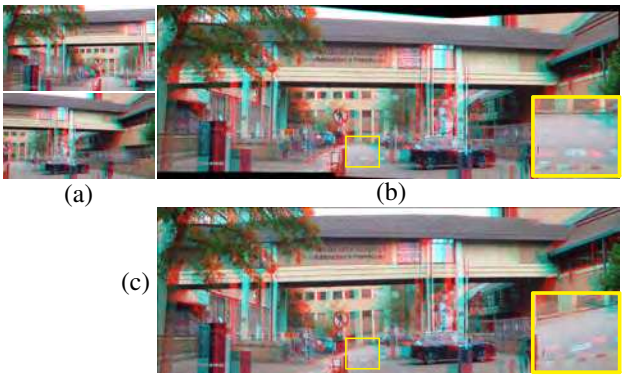


Fig. 11 A failure case with large distortion. (a) input image pair, (b) result of global stitching, (c) our final stitching result with rectangular boundary.

7 CONCLUSION

We have proposed a new stereoscopic image stitching method with rectangular boundaries, which improves the visual quality of the stitching results for the final display by providing rectangular boundaries. We propose a global optimization to stitch left and right views, which ensures the disparity consistency and boundary regularity. In our optimization framework, we first simultaneously consider feature alignment, disparity consistency, shape preserving and global similarity constraint. Then we extract the boundaries of panoramic image pair by polygon Boolean operations and reconstruct the target regular boundary. The second global optimization incorporating the rectangular boundary constraint and line preserving constraint is performed to warp and blend images for the final stereoscopic panorama. To avoid undesirable distortion introduced by content missing in regions with salient structures, we further propose content-aware rectangling, which can progressively perform rectangling of the panorama with irregular boundary.

In the future, we will further study stereoscopic video stitching, and there might be more challenges due to the free movement of cameras. We need to take measures to ensure the seamless stitching with unnoticeable shak-

ing in a unified optimization [9,32]. In addition, we also need to consider the visual comfort [19,6] in stereoscopic video stitching for better viewing experiences.

Acknowledgements This work was supported by the National Natural Science Foundation of China (Grant No. 61602402), Zhejiang Province Public Welfare Technology Application Research (Grant No. LGG19F020001), and the Royal Society (Grant No. IES\R1\180126).

Compliance with Ethical Standards

Conflict of interest The authors declare that they have no conflict of interest.

References

1. Brown, M., Lowe, D.G.: Automatic panoramic image stitching using invariant features. International Journal of Computer Vision **74**(1), 59–73 (2007)
2. Chang, C., Liang, C., Chuang, Y.: Content-aware display adaptation and interactive editing for stereoscopic images. IEEE Trans. Multimedia **13**(4), 589–601 (2011)
3. Chang, C., Sato, Y., Chuang, Y.: Shape-preserving half-projective warps for image stitching. In: IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp. 3254–3261 (2014)
4. Chen, Y., Chuang, Y.: Natural image stitching with the global similarity prior. In: Proceedings of the European Conference on Computer Vision (ECCV), pp. 186–201 (2016)
5. Du, S., Hu, S., Martin, R.R.: Changing perspective in stereoscopic images. IEEE Trans. Vis. Comput. Graph. **19**(8), 1288–1297 (2013)
6. Du, S., Masiá, B., Hu, S., Gutierrez, D.: A metric of visual comfort for stereoscopic motion. ACM Trans. Graph. **32**(6), 222:1–222:9 (2013)
7. Gao, J., Kim, S.J., Brown, M.S.: Constructing image panoramas using dual-homography warping. In: IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp. 49–56 (2011)
8. von Gioi, R.G., Jakubowicz, J., Morel, J., Randall, G.: LSD: A fast line segment detector with a false detection control. IEEE Trans. Pattern Anal. Mach. Intell. **32**(4), 722–732 (2010)
9. Guo, H., Liu, S., He, T., Zhu, S., Zeng, B., Gabbouj, M.: Joint video stitching and stabilization from moving cameras. IEEE Trans. Image Processing **25**(11), 5491–5503 (2016)
10. He, K., Chang, H., Sun, J.: Rectangling panoramic images via warping. ACM Trans. Graph. **32**(4), 79:1–79:10 (2013)
11. Lang, M., Hornung, A., Wang, O., Poulakos, S., Smolic, A., Gross, M.H.: Nonlinear disparity mapping for stereoscopic 3d. ACM Trans. Graph. **29**(4), 75:1–75:10 (2010)
12. Lee, S., Kim, Y., Lee, J., Kim, K., Lee, K., Noh, J.: Depth manipulation using disparity histogram analysis for stereoscopic 3d. The Visual Computer **30**(4), 455–465 (2014)
13. Lin, C.C., Pankanti, S.U., Ramamurthy, K.N., Aravkin, A.Y.: Adaptive as-natural-as-possible image stitching. In: IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp. 1155–1163 (2015)
14. Lin, K., Jiang, N., Cheong, L., Do, M.N., Lu, J.: SEAGULL: seam-guided local alignment for parallax-tolerant image stitching. In: Proceedings of the European Conference on Computer Vision (ECCV), pp. 370–385 (2016)
15. Lin, K., Liu, S., Cheong, L., Zeng, B.: Seamless video stitching from hand-held camera inputs. Comput. Graph. Forum **35**(2), 479–487 (2016)
16. Liu, S., Yuan, L., Tan, P., Sun, J.: Bundled camera paths for video stabilization. ACM Trans. Graph. **32**(4), 78:1–78:10 (2013)
17. Liu, Y., Sun, L., Yang, S.: A retargeting method for stereoscopic 3d video. Computational Visual Media **1**(2), 119–127 (2015)
18. Martínez, F., Rueda, A.J., Feito, F.R.: A new algorithm for computing boolean operations on polygons. Computers & Geosciences **35**(6), 1177–1185 (2009)
19. Mu, T., Sun, J., Martin, R.R., Hu, S.: A response time model for abrupt changes in binocular disparity. The Visual Computer **31**(5), 675–687 (2015)
20. Mu, T., Wang, J., Du, S., Hu, S.: Stereoscopic image completion and depth recovery. The Visual Computer **30**(6-8), 833–843 (2014)
21. Szeliski, R.: Image alignment and stitching: A tutorial. Foundations and Trends in Computer Graphics and Vision **2**(1) (2006)
22. Tang, Y., Tong, R., Tang, M., Zhang, Y.: Depth incorporating with color improves salient object detection. The Visual Computer **32**(1), 111–121 (2016)
23. Tong, R., Zhang, Y., Cheng, K.: Stereopasting: Interactive composition in stereoscopic images. IEEE Trans. Vis. Comput. Graph. **19**(8), 1375–1385 (2013)
24. Wang, H., Zhou, Y., Wang, X., Fang, L.: A natural shape-preserving stereoscopic image stitching. In: IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), pp. 1812–1816 (2018)
25. Wang, M., Zhang, X., Liang, J., Zhang, S., Martin, R.R.: Comfort-driven disparity adjustment for stereoscopic video. Computational Visual Media **2**(1), 3–17 (2016)
26. Yan, W., Hou, C., Lei, J., Fang, Y., Gu, Z., Ling, N.: Stereoscopic image stitching based on a hybrid warping model. IEEE Trans. Circuits Syst. Video Techn. **27**(9), 1934–1946 (2017)
27. Yan, W., Hou, C., Wang, B., Wang, L.: Content-aware disparity adjustment for different stereo displays. Multimedia Tools Appl. **76**(8), 10,465–10,479 (2017)
28. Zaragoza, J., Chin, T., Tran, Q., Brown, M.S., Suter, D.: As-projective-as-possible image stitching with moving DLT. IEEE Trans. Pattern Anal. Mach. Intell. **36**(7), 1285–1298 (2014)
29. Zhang, F., Liu, F.: Parallax-tolerant image stitching. In: IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp. 3262–3269 (2014)
30. Zhang, F., Liu, F.: Casual stereoscopic panorama stitching. In: IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp. 2002–2010 (2015)
31. Zhang, Y., Lai, Y., Zhang, F.: Content-preserving image stitching with regular boundary constraints. ArXiv e-prints (2018)
32. Zhu, Z., Lu, J., Wang, M., Zhang, S., Martin, R.R., Liu, H., Hu, S.: A comparative study of algorithms for realtime panoramic video blending. IEEE Trans. Image Processing **27**(6), 2952–2965 (2018)